GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 09:38:13 ; Search time 2727.16 Seconds
(without alignments)
-3173.152 Million cell updates/sec

Title: US-09-209-069-17
Perfect score: 2850
Sequence: 1 tccatcctaatacgactcac.....aaaataatgtgggggtgtgg 2850

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 821193 seqs, -1518192014 residues

Total number of hits satisfying chosen parameters: 1642386

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2715.4	95.3	2802	11	AF084644	AF084644 Homo sapi
	2	2586.2	90.7	2905	11	AF084645	AF084645 Homo sapi
	3	2576.8	90.4	4448	10	HSAJ9936	AJ009936 Homo sapi
	4	2344.8	82.3	4337	10	HSAJ9937	AJ009937 Homo sapi
	5	1866.6	65.5	2146	40	AF061056	AF061056 Homo sapi
	6	1046	36.7	1895	3	AF182217	AF182217 Oryctolag
	7	974.2	34.2	1709	12	AF031814	AF031814 Mus muscu
	8	967.8	34.0	1755	12	AF151377	AF151377 Rattus no
c	9	343	12.0	365	13	G36929	G36929 SHGC-56597
	10	280.8	9.9	1377	12	MUSVDR	D31969 Mouse mRNA
	11	278.8	9.8	2043	12	RATDHVRD3	J04147 Rat 1,25-di
	12	275	9.6	2494	4	AF011356	AF011356 Gallus ga
	13	272.8	9.6	1284	11	AF026260	AF026260 Homo sapi
	14	272.8	9.6	1335	9	HSVD3R	X67482 H.sapiens m
	15	272.8	9.6	4604	9	HUMVDR	J03258 Human vitam
	16	267.8	9.4	1724	4	CJU12641	U12641 Coturnix ja
	17	258.6	9.1	1638	4	XLRNAONST	X75163 X.laevis mR
	18	236	8.3	1071	5	E14584	E14584 Rat mRNA is
	19	232	8.1	1782	4	XLU91846	U91846 Xenopus lae
	20	188.6	6.6	1404	5	E14585	E14585 Human mRNA
	21	172.2	6.0	1450	9	HSONHORE	Z30425 H.sapiens m
	22	170.6	6.0	1450	5	AR009748	AR009748 Sequence
	23	170.6	6.0	1450	5	I73479	I73479 Sequence 1
	24	170.6	6.0	1450	5	I81188	I81188 Sequence 1
	25	168.6	5.9	1350	12	AF009327	AF009327 Mus muscu
	26	117.4	4.1	1332	12	AF009328	AF009328 Mus muscu
	27	100.8	3.5	2797	35	DMU36792	U36792 Drosophila
	28	100.8	3.5	113474	43	AC013932	AC013932 Drosophil
c	29	100.8	3.5	116280	33	AC007853	AC007853 Drosophil
	30	100.8	3.5	132103	33	AC008206	AC008206 Drosophil
	31	100.4	3.5	203	9	HUMVDRPBMC	M65208 Human vitam
	32	97.6	3.4	1898	5	I46765	I46765 Sequence 1
	33	97.6	3.4	1979	5	AR035536	AR035536 Sequence
	34	97.6	3.4	1979	24	E11456	E11456 cDNA encodi
	35	97.6	3.4	2010	9	HSU07132	U07132 Human stero
	36	97.6	3.4	2030	5	I36667	I36667 Sequence 1
	37	97.6	3.4	2030	5	I70211	I70211 Sequence 1
	38	97.2	3.4	1746	9	HSU14534	U14534 Human orpha
	39	94.4	3.3	1494	4	GGRARA21	X78335 G.gallus re
	40	94.4	3.3	1552	4	GGRARA1	X73972 G.gallus RA
	41	93	3.3	2908	5	I08117	I08117 Sequence 1
	42	91.4	3.2	1284	10	S50916	S50916 PML-RAR fus
	43	91.4	3.2	1920	9	HSRAR	X06538 Human mRNA
	44	91.4	3.2	1944	10	HSU41743	U41743 Human nucle
	45	91.4	3.2	2073	10	HSU41742	U41742 Human nucle

ALIGNMENTS

RESULT 1
AF084644
LOCUS AF084644 2802 bp mRNA PRI 20-OCT-1998
DEFINITION Homo sapiens orphan nuclear receptor (PAR2) mRNA, complete cds.
ACCESSION AF084644
VERSION AF084644.1 GI:3769536
KEYWORDS .

GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 29, 2000, 10:21:30 ; Search time 22.21 Seconds
(without alignments)
504.436 Million cell updates/sec

Title: US-09-209-069-18
Perfect score: 2494
Sequence: 1 MTVTRTHHFKEGSLRAPAIP.....QDIHPFATPLMQELFGITGS 473

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	979.5	39.3	386	1	R98521	Xenopus orphan rec
2	801	32.1	423	1	W94623	Rat vitamin D rece
3	798	32.0	423	1	W47509	Rat vitamin D rece
4	791	31.7	477	1	Y09035	Human vitamin D re
5	789	31.6	427	1	W68156	Human vitamin D re
6	789	31.6	427	1	Y09064	Human vitamin D re
7	789	31.6	450	1	Y09036	Human vitamin D re
8	768	30.8	348	1	W32536	Constitutively act
9	768	30.8	348	1	W93902	Human CAR receptor
10	763	30.6	348	1	R41346	Human CAR receptor
11	725	29.1	358	1	W93903	Mouse CAR receptor
12	656	26.3	356	1	W37261	Rat vitamin D rece
13	656	26.3	356	1	W94622	Rat vitamin D rece
14	632	25.3	367	1	R43656	Rat vitamin D rece
15	500.5	20.1	460	1	R74738	Human ubiquitous n
16	500.5	20.1	461	1	R52980	Human recombinant
17	500.5	20.1	461	1	R98140	NER receptor poten
18	500.5	20.1	461	1	R97982	Human steroid rece
19	498	20.0	460	1	W25034	Human ubiquitous n
20	497.5	19.9	461	1	R96234	Human foetal lung
21	491.5	19.7	447	1	W03326	LXR-alpha, orphan
22	490	19.6	446	1	R99736	Retinoid X recepto
23	488	19.6	446	1	R94169	OR-1 orphan recept
24	483.5	19.4	443	1	R74739	Rat ubiquitous nuc
25	483.5	19.4	443	1	W25035	Rat ubiquitous nuc
26	481.5	19.3	440	1	R33744	XR2. DNA encoding
27	472.5	18.9	472	1	W40072	Human retinoid rec
28	461	18.5	469	1	W03448	Farnesoid-activate
29	455.5	18.3	757	1	W71297	Lucilia cuprina ec
30	452.5	18.1	484	1	R99735	Retinoid X recepto
31	450.5	18.1	456	1	P80921	Sequence encoded b
32	445	17.8	451	1	R99739	Retinoid X recepto
33	439.5	17.6	579	1	R71565	EAR-1r gene produc
34	434.5	17.4	878	1	R13793	Ecdysone receptor.
35	432	17.3	514	1	R47621	Rat thyroid hormon
36	432	17.3	746	1	W33655	Modified ecdysone
37	431.5	17.3	878	1	R32889	DHR23alpha protein
38	428	17.2	448	1	P90341	hap (hepatoma) pro
39	427	17.1	448	1	R10548	Murine Retinoic Ac
40	427	17.1	448	1	R84725	Murine retinoic ac
41	426	17.1	448	1	R55128	Retinoic acid rece
42	425	17.0	448	1	R66764	Human RAR-beta. De
43	425	17.0	448	1	R85315	Human hepatoma ret
44	418	16.8	410	1	R78318	Human thyroid horm
45	417	16.7	606	1	W95701	Bombyx mori nuclea

ALIGNMENTS

RESULT 1
R98521
ID R98521 standard; Protein; 386 AA.
AC R98521;
DT 14-NOV-1996 (first entry)
DE Xenopus orphan receptor 6.
KW Xenopus orphan receptor 6; XOR-6; steroid receptor; vitamin D;
KW hydroxybenzoate; mercaptobenzoate; aminobenzoate.
OS Xenopus laevis.
FH Key Location/Qualifiers
FT domain 37. .102
FT /label= DNA_binding_domain
FT domain 183. .386
FT /label= Ligand_binding_domain
PN WO9622390-A1.
PD 25-JUL-1996.
PF 16-JAN-1996; U00058.
PR 17-JAN-1995; US-374445.
PA (SALK) SALK INST BIOLOGICAL STUDIES.
PI Blumberg B, Evans RM, Umesono K;
DR WPI; 96-354546/35.
DR N-PSDB; T36499.
PT DNA encoding receptor polypeptide responsive to hydroxy, mercapto or
PT amino benzoate(s) - useful to regulate gene transcription
PS Claim 4; Page 27-28; 42pp; English.
CC Xenopus orphan receptor 6 (XOR-6) (R98521) is a new member of the
CC steroid receptor superfamily, characterised as being responsive to
CC the presence of hydroxy, mercapto or amino benzoate(s) and as
CC regulating the transcription of associated gene(s). It shows 73%
CC identity in the DNA binding domain, and 42% identity in the ligand
CC binding domain, to the human vitamin D receptor. Recombinant XOR-6
CC can be expressed in animal cells; a cDNA clone (T36499) coding for
CC XOR-6 has been isolated. The recombinant XOR-6 may be used to
CC regulate gene transcription or to raise antibodies of diagnostic
CC or therapeutic appln.
SQ Sequence 386 AA;

Query Match 39.3%; Score 979.5; DB 1; Length 386;
Best Local Similarity 48.6%; Pred. No. 2e-90;
Matches 202; Conservative 60; Mismatches 105; Indels 49; Gaps 9;

```
Qy      57 EDTESVPGKPSVNADEEVGGPQICRVCGDKATGYHFNVMTCEGCKGFFRRAMKRNARLRC 116
           |: |          ||: | |:||| |||:||||||| ||||||||||||||:||| || |
Db      14 EEEEDASNSCGTGEDEDDGDPKICRACGDRATGYHFNAMTCEGCKGFFRRAVKRNLRLSC 73

Qy     117 PFRKGACEITRKTRRQCQACRLRKCLESGMKKEMIMSDEAVEERRALIKRK-KSERTGTQ 175
           || : :| | :  || |||||:|||: ||:||:|||| |||:||||||||  |  :
Db      74 PF-QNSCVINKSNRRHCQACRLKKCLDIGMRKELIMSDAAVEQRRALIKRKHKLTKLPPT 132

Qy     176 PLGVQGLTEEQRMMIRELMDAQMKTFDTTFSHFKNFRLPGVLSSGCELPESLQAPSREEA 235
           | |  || ||:  : :|: |  |||| |:  |||| |                  | |
Db     133 PPGA-SLTPEQQHFLTQLVGAHTKTFDFNFTFSKNFR----------------PIR--- 172

Qy     236 AKWSQVRKDLCSLKVSLQLRGEDGSVWNYKPPADSGGKEIFSLLPHMADMSTYMFKGIIS 295
                               |  | :    :| | |   | | :|||::|: ||| |||||
Db     172 ------------------RSSDPT---QEPQATS--SEAFLMLPHISDLVTYMIKGIIS 207
```

```
OM of: US-09-209-069-18 to: GenEmbl:*   out_format : pfs

Date: Mar 29, 2000 10:59 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09209069/runat_28032000_145604_25682/app_query.fasta.1
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US09209069
 -NCPU=6 -ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-209-069-18
Query length: 473
Database: GenEmbl:*
Database sequences: 821193
Database length: -1518192014
Search time (sec): 1139.270000


score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pr3:AF084644        +  2494.00 3680.74  5.4e-197  2802  ! AF084644 Homo sapiens orphan n
gb_pr4:AF061056        +  2334.00 3445.96  6.4e-184  2146  ! AF061056 Homo sapiens orphan n
gb_pr3:AF084645        +  2334.00 3442.90  9.5e-184  2905  ! AF084645 Homo sapiens orphan n
gb_pr2:HSAJ9936        +  2330.00 3432.66  3.5e-183  4448  ! AJ009936 Homo sapiens mRNA for
gb_pr2:HSAJ9937        +  2110.50 3107.14  4.8e-165  4337  ! AJ009937 Homo sapiens mRNA for
gb_om:AF182217         +  1888.50 2786.01  3.7e-147  1895  ! AF182217 Oryctolagus cuniculus
gb_ro:AF031814         +  1796.50 2650.51  1.3e-139  1709  ! AF031814 Mus musculus pregnane
gb_ro:AF151377         +  1785.50 2633.92  1.1e-138  1755  ! AF151377 Rattus norvegicus pre
gb_ov:XLRNAONST        +   992.00 1456.90  4.0e-73   1638  ! X75163 X.laevis mRNA for orpha
gb_ov:CJU12641         +   815.50 1194.43  1.7e-58   1724  ! U12641 Coturnix japonica vitam
gb_ov:AF011356         +   815.50 1190.69  2.7e-58   2494  ! AF011356 Gallus gallus vitamin
gb_ov:XLU91846         +   814.00 1191.87  2.3e-58   1782  ! U91846 Xenopus laevis vitamin
gb_ro:RATDHVRD3        +   807.50 1180.84  9.5e-58   2043  ! J04147 Rat 1,25-dihydroxyvitam
gb_pr1:HUMVDR          +   795.00 1154.07  2.9e-56   4604  ! J03258 Human vitamin D recepto
gb_ro:MUSVDR           +   792.50 1162.56  9.9e-57   1377  ! D31969 Mouse mRNA for vitamin
gb_pr3:AF026260        +   789.00 1158.07  1.8e-56   1284  ! AF026260 Homo sapiens vitamin
gb_pr1:HSVD3R          +   789.00 1157.68  1.8e-56   1335  ! X67482 H.sapiens mRNA for 1,25
gb_pat:AR009748        +   783.00 1147.94  6.4e-56   1450  ! AR009748 Sequence 1 from paten
gb_pat:I73479          +   783.00 1147.94  6.4e-56   1450  ! I73479 Sequence 1 from patent
gb_pat:I81188          +   783.00 1147.94  6.4e-56   1450  ! I81188 Sequence 1 from patent
gb_pr1:HSONHORE        +   783.00 1147.94  6.4e-56   1450  ! Z30425 H.sapiens mRNA for orph
gb_ro:AF009327         +   725.00 1062.58  3.7e-51   1350  ! AF009327 Mus musculus orphan n
gb_pat:E14584          +   656.00  962.51  1.4e-45   1071  ! E14584 Rat mRNA isoform for Vi
gb_ro:AF009328         +   613.50  897.23  5.9e-42   1332  ! AF009328 Mus musculus orphan n
gb_pat:E14585          +   610.50  892.24  1.1e-41   1404  ! E14585 Human mRNA isoform for
gb_ro:MMU132601        +   504.50  733.07  8.3e-33   1686  ! AJ132601 Mus musculus mRNA for
gb_pr1:HSU14534        +   503.50  731.23  1.0e-32   1746  ! U14534 Human orphan receptor m
gb_pat:AR035536        +   503.50  729.96  1.2e-32   1979  ! AR035536 Sequence 1 from paten
em_pat:E11456          +   503.50  729.96  1.2e-32   1979  ! E11456 cDNA encoding a steroid
gb_pr1:HSU07132        +   503.50  729.81  1.3e-32   2010  ! U07132 Human steroid hormone r
gb_pat:I36667          +   503.50  729.71  1.3e-32   2030  ! I36667 Sequence 1 from patent
gb_pat:I70211          +   503.50  729.71  1.3e-32   2030  ! I70211 Sequence 1 from patent
gb_pat:I46765          +   500.50  725.93  2.1e-32   1898  ! I46765 Sequence 1 from patent
gb_ro:AF085745         +   498.50  724.29  2.5e-32   1665  ! AF085745 Mus musculus nuclear
gb_ro:RNU11685         +   496.50  720.97  3.9e-32   1723  ! U11685 Rattus norvegicus orpha
gb_ro:MMU09419         +   493.50  715.85  7.5e-32   1841  ! U09419 Mus musculus retinoid X
gb_pr2:HSU22662        +   491.50  714.77  8.6e-32   1528  ! U22662 Human nuclear orphan re
gb_pat:I28453          +   491.50  713.93  9.6e-32   1659  ! I28453 Sequence 7 from patent
gb_pat:I81148          +   491.50  713.93  9.6e-32   1659  ! I81148 Sequence 7 from patent
gb_pat:I84701          +   491.50  713.93  9.6e-32   1659  ! I84701 Sequence 7 from patent
gb_ro:RNU20389         +   488.00  707.49  2.2e-31   1877  ! U20389 Rattus norvegicus orpha
gb_pat:A49139          +   488.00  707.19  2.3e-31   1934  ! A49139 Sequence 1 from Patent
gb_pat:I46766          +   483.50  700.38  5.5e-31   1959  ! I46766 Sequence 3 from pate
gb_ro:RNU14533         +   483.50  700.38  5.5e-31   1959  ! U14533 Rattus norvegicus or
gb_pr2:HSU68233        +   472.50  682.80  5.2e-30   2218  ! U68233 Human farnesol recep

seq_name: gb_pr3:AF084644

seq_documentation_block:
LOCUS       AF084644     2802 bp    mRNA            PRI       20-OCT-1998
DEFINITION  Homo sapiens orphan nuclear receptor (PAR2) mRNA, complete cds.
ACCESSION   AF084644
VERSION     AF084644.1  GI:3769536
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2802)
  AUTHORS   Bertilsson,G., Heidrich,J., Svensson,K., Asman,M., Jendeberg,L.,
            Sydow-Backman,M., Ohlsson,R., Postlind,H., Blomquist,P. and
            Berkenstam,A.
  TITLE     Identification of a human nuclear receptor defines a new signaling
            pathway for CYP3A induction
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12208-12213 (1998)
  MEDLINE   98445350
REFERENCE   2  (bases 1 to 2802)
  AUTHORS   Bertilsson,G., Asman,M., Blomquist,P. and Berkenstam,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-AUG-1998) Cell and Molecular Biology, Medical Nobel
            Institute, Karolinska Institute, Doktorsringen2, Stockholm 17177,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .2802
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="liver"
     gene            1. .2802
                     /gene="PAR2"
     CDS             60. .1481
                     /gene="PAR2"
                     /codon_start=1
                     /product="orphan nuclear receptor"
                     /protein_id="AAC64557.1"
                     /db_xref="GI:3769537"
                     /translation="MTVTRTHHFKEGSLRAPAIPLHSAAAELASNHPRGPEANLEVRP
                     KESWNHADFVHCEDTESVPGKPSVNADEEVGGPQICRVCGDKATGYHFNVMTCEGCKG
                     FFRRAMKRNARLRCPFRKGACEITRKTRRQCQACRLRKCLESGMKKEMIMSDEAVEER
                     RALIKRKKSERTGTQPLGVQGLTEEQRMMIRELMDAQMKTFDTTFSHFKNFRLPGVLS
                     SGCELPESLQAPSREEAAKWSQVRKDLCSLKVSLQLRGEDGSVWNYKPPADSGGKEIF
                     SLLPHMADMSTYMFKGIISFAKVISYFRDLPIEDQISLLKGAAFELCQLRFNTVFNAE
                     TGTWECGRLSYCLEDTAGGFQQLLLEPMLKFHYMLKKLQLHEEEYVLMQAISLFSPDR
                     PGVLQHRVVDQLQEQFAITLKSYIECNRPQPAHRFLFLKIMAMLTELRSINAQHTQRL
                     LRIQDIHPFATPLMQELFGITGS"
BASE COUNT      723 a    715 c    755 g    609 t
ORIGIN


alignment_scores:
           Quality: 2494.00              Length:    473
             Ratio:  5.273                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-209-069-18 x AF084644   ..

 Align seg 1/1  to: AF084644  from: 1  to: 2802

       1 MetThrValThrArgThrHisHisPheLysGluGlySerLeuArgAlaPr 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      60 ATGACAGTCACCAGGACTCACCACTTCAAGGAGGGGTCCCTCAGAGCACC 109

      17 oAlaIleProLeuHisSerAlaAlaAlaGluLeuAlaSerAsnHisProA 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     110 TGCCATACCCCTGCACAGTGCTGCGGCTGAGTTGGCTTCAAACCATCCAA 159
```